

## IDENTIFICAÇÃO DE REGIÕES PLUVIOMETRICAMENTE HOMOGÊNEAS NO NORDESTE DO BRASIL USANDO ANÁLISE MULTIVARIADA

SANTOS ARAÚJO, Winícius - winicius@dca.ufcg.edu.br

Dep Ciências Atmosféricas - DCA/Universidade Federal de Campina Grande – UFCG

SAVIANO DE SOUSA, Francisco de A. - fassis@dca.ufcg.edu.br

Dep Ciências Atmosféricas - DCA/Universidade Federal de Campina Grande – UFCG

---

**RESUMO.** São aplicadas técnicas multivariadas às séries de precipitação pluvial na região Nordeste do Brasil. Utilizaram-se séries de dados mensais de precipitação para todo o Nordeste brasileiro oriundas das séries do National Climatic Data Center (NCDC) gradeadas numa resolução espacial de  $0,5^\circ \times 0,5^\circ$ , para o período de 1961-1990. Usaram-se os métodos de Componentes Principais e Análise de Agrupamentos para determinar os padrões espacial e temporal e regiões homogêneas com relação à precipitação pluvial. Os resultados obtidos mostraram que quatro regiões climaticamente homogêneas podem caracterizar a distribuição da precipitação no Nordeste do Brasil, quais sejam: a identificada como grupo 1 que representa a região do norte (meio-norte) e tem como principal sistema indutor de chuva os vórtices ciclônicos de ar superior; a identificada como grupo 2, que engloba o semiárido nordestino e está sob influência da ação direta da zona de convergência intertropical; o grupo 3, cuja proximidade com o cerrado central se faz presente e é influenciado pelas frentes frias provenientes da parte mais ao sul do país e, finalmente o grupo 4, o qual predominantemente é costeiro, se estendendo deste do sul da Bahia até o extremo leste do Rio Grande do Norte e tem como principal sistema indutor de chuva as ondas de leste, e em menor escala o efeito das brisas.

**Palavras-chave:** nordeste, precipitação, homogeneidade, fatores, agrupamento.

---

### *RAINFALL HOMOGENEOUS REGIONS IDENTIFICATION IN NORTHEASTERN BRAZIL USING MULTIVARIATE ANALYSIS*

**ABSTRACT.** Applied techniques multivariate to the pluvial precipitation series in the Northeast region of Brazil. Series of monthly precipitation data had been used for the Brazilian Northeast deriving of the series of National Climatic Data Center (NCDC) grads in a space resolution of  $0.5^\circ \times 0.5^\circ$ , for the period of 1961-1990. The methods of Principal Components and Cluster Analysis had been used to determine the standards spaces and time and homogeneous regions with regard to the pluvial precipitation. The gotten results had shown that four climatic homogeneous regions can characterize the distribution of the Northeast precipitation of Brazil, which are: the identified one as group 1 that it more represents the region to the north (half-north) and has as main inductive rain system the cyclone vortices of superior air; the identified one as group 2, that extent the semi-arid northeastern and is under influence of the direct action of the zone of intertropical convergence; group 3, whose proximity with the central pasture if makes present and is influenced by the cold fronts comings of the part more to the south of the country and, finally group 4, which predominantly is coastal, if extending of this of the south of the Bahia until the extreme east of the Rio Grande do Norte and has as main inductive rain system the east waves, and in lesser the effect of breezes scales.

**Key-words:** northeast, rainfall, homogeneity, factors, cluster

---

## INTRODUÇÃO

Dentre inúmeras pesquisas que abordam a variabilidade climática, tem-se destacado a identificação de fenômenos atmosféricos associados aos ciclos da variabilidade natural, cuja importância reside no provimento de subsídios para formulação de modelos numéricos que descrevam a distribuição espacial de séries temporais de regiões que sejam climaticamente homogêneas, especialmente aquelas de precipitação, cuja influência atinge diretamente a vida no Nordeste brasileiro, sobretudo perante possíveis alterações climáticas regionais atuais.

Visando classificar regiões homogêneas quanto à uma determinada variável, a Análise de Agrupamento (AA) tem se consolidado como uma das maneiras mais úteis no processo de filtrar dados para descobrir grupos homogêneos e identificar padrões ocultos nos dados (HALKIDI *et al.*, 2001), embora seja possível efetuar de forma relativamente simples e rápida a identificação de áreas homogêneas, tomando por base planos fatoriais (ACP - Análise de Componentes Principais).

O objetivo deste trabalho é aplicar a ACP para obter a distribuição espaço-temporal da precipitação e a AA para estabelecer a regionalização da homogeneidade. Portanto, apresentar a robustez da técnica espaço-temporal, mostrando um estudo da variação espacial e temporal

da precipitação do Nordeste do Brasil usando análise de componentes principais e seu estabelecimento de grupos homogêneos de precipitação por meio da análise de agrupamento para o período de 1961-1990.

### **Análise Multivariada Subjetiva - Análise de Componentes Principais (ACP)**

A Análise de componentes principais (ACP) é uma técnica multivariada usada para transformar um conjunto de variáveis relacionadas (ou correlacionadas) em um conjunto de variáveis não-relacionadas (ou não-correlacionadas) que quantificam em proporções decrescentes a variação dos dados originais. A razão de utilização deste método é uma tentativa de reduzir a complexidade existente nos dados para um número reduzido de variáveis que necessariamente devam ser consideradas.

Originalmente a Análise de Componentes Principais (ACP) foi introduzida por Pearson em 1901 e por Hotelling em 1933, Everitt e Der (1977). A aplicação da ACP a um conjunto de dados de grande dimensão é interessante, inicialmente, apenas para determinar combinações lineares das variáveis originais que expliquem o máximo possível a variação existente nos dados iniciais, Silva (2001). A ACP transforma um sistema de  $p$  eixos originais em um novo sistema que tem duas importantes propriedades: os novos eixos principais são ortogonais e as componentes principais são arranjadas em ordem decrescente de importância, de modo que cada uma delas explique a máxima quantidade da variância das observações, não-explicada pelas componentes anteriores, Assis *et al.* (1996).

SILVA (1996) utilizou dados de chuva coletados em 68 postos pluviométricos distribuídos no Estado da Paraíba no período 1930-1993. O objetivo do autor na aplicação das técnicas de análise em componentes principais foi identificar sub-regiões pluviometricamente homogêneas e determinar as mudanças que o fenômeno ENOS provoca na distribuição espacial da precipitação no Estado da Paraíba. A aplicação da análise de componentes principais para o conjunto de dados mostrou que as duas primeiras componentes explicam 96% da variância espacial total. O autor concluiu que o conjunto de dados evidenciou diferenças significativas na distribuição espacial da chuva no Estado da Paraíba entre anos de ENOS fortes e anos climatológicos.

Prates (1994) aplicou a metodologia da análise de componentes principais em dados de chuva coletados em 128 postos pluviométricos distribuídos no estado de Minas Gerais e regiões limítrofes de Goiás, Espírito Santo e Bahia. O objetivo do autor foi identificar regiões pluviometricamente homogêneas. Foram analisadas as três primeiras componentes principais que explicaram 92,32% da variância total dos dados. O padrão espacial da primeira componente principal, que explicou 86,85 da variância, foi associado à resposta local a mecanismos de grande escala. Os padrões espaciais da segunda e terceira componentes principais, que explicaram 3,55% e 1,92% da variância, respectivamente, podem representar diferenças no regime pluviométrico associadas a efeitos locais, apesar da pequena porcentagem de explicação da variância.

Maldonado *et al.* (2000) utilizaram a análise de componentes principais para verificar as mudanças na cobertura vegetal em uma mesma região semiárida em Pernambuco, Brasil. Foram utilizadas duas imagens da faixa TM-3 em duas datas: 1984 e 1996. A segunda componente principal foi dividida em cinco classes, com os limites localizados em 1 e 2 desvios padrões a partir da média. O método se mostrou adequado para a detecção das mudanças na cobertura vegetal e no uso da terra na área da Caatinga estudada.

Andrade *et al.* (2003) identificaram e avaliaram a importância das variáveis envolvidas no processo da evapotranspiração pela análise de componentes principais. Essa técnica foi aplicada com o propósito de identificar a importância relativa das diferentes variáveis envolvidas na evapotranspiração em bacias hidrográficas da Região Centro Sul do Estado do Ceará. A técnica da análise de componentes principais demonstrou que um modelo com duas componentes é adequado para representar o processo de evapotranspiração na área estudada. Os resultados revelaram que as variáveis climatológicas com maiores representatividades no processo de evapotranspiração foram: razão de insolação, velocidade do vento e precipitação; enquanto a temperatura máxima e mínima do ar apresentou menor influência para a região do Centro Sul do Estado do Ceará.

Diniz (2007) utilizando a técnica estatística multivariada de Componentes Principais para estudar as cheias na Bacia hidrográfica do São Francisco em abril de 1985 a janeiro de 2004

afim de identificar variáveis meteorológicas dominantes no processo de formação e evolução de SC (Sistemas Convectivos) intensos que contribuíram para os episódios de chuvas extremas e ocorrência de cheia, verificou que os efeitos locais foram significativos para a evolução dos SC intensos observados em abril de 1985 e, utilizando dados de reanálises (NCEP - National Centers for Environment Prediction) verificou que um modelo com 4 componentes foi adequado para representar a estrutura inicial das variáveis.

Silva (2004), com base em dados da Região Nordeste do Brasil e utilizando ACP, desenvolveu um estudo com objetivo de identificar as variáveis meteorológicas que apresentam maior influência no processo de ET (evapotranspiração) na região, e verificou que através da referida técnica aplicada a 88 estações meteorológicas distribuídas pelo NEB, revelou que o processo de evapotranspiração para a área estudada é governado basicamente por duas componentes, sendo a primeira o gradiente de pressão de vapor do ar e a segunda a da temperatura do ar.

### **Análise Multivariada Objetiva - Análise de Agrupamento (AA) (Cluster Analysis)**

A AA tem obtido reconhecimento como uma efetiva técnica estatística para lidar com agrupamento de estações dentro de regiões climatologicamente homogêneas (DE GAETANO e SHULMAN, 1990; AHMED, 1997; DE GAETANO, 2001), ou para agrupar períodos temporais dentro de grupos que reflitam a ocorrência de eventos de tempo ou padrões (RAMOS, 2001). Além de cada grupo fornecer uma maneira conveniente de resumir os dados climáticos concisamente (DE GAETANO, 2001) é uma técnica isenta de subjetividade e justificada em critérios estatísticos para realizar a identificação final de zonas homogêneas de precipitação ou de outra variável identificada.

A Análise de agrupamento é aplicada à pesquisa geofísica desde 1961. Foi durante esta década que esta técnica se proliferou. Uma pesquisa detalhada da literatura de análise de agrupamento usada nos campos da geofísica foi realizada dando conta que vários métodos de análise de agrupamento foram extensamente usados em tipos diferentes de problemas em pesquisa atmosféricas (GONG e RICHMAN, 1995).

Silva *et al.* (1996) objetivando estudar o regime pluviométrico regional e temporal de algumas eco-regiões costeiras no Nordeste do Brasil, especialmente em áreas de planalto e baixadas litorâneas, através de análise dos dados feita considerando médias mensais de chuva coletadas em 127 estações climatológicas, obteve quatro sub-regiões com características homogêneas, localizadas ao longo desta região. Na análise de cluster usou-se o método de Ward para avaliar a chuva associada à climatologia da região, para assim classificar a região quanto a homogeneidade.

Braga *et al.* (1998) objetivando dar uma contribuição para o estudo climático ao estado da Bahia, no nordeste brasileiro, identificou sub-regiões homogêneas em relação ao ciclo anual e totais médios de precipitação. As regiões foram identificadas usando o método de análise de cluster proposto por Ward (1963) e os resultados obtidos mostraram que o estado da Bahia pode ser subdividido em 9 áreas de regimes homogêneos de precipitação.

Santos *et al.* (2008) buscando ampliar a difusão do método AA entre os trabalhos envolvendo a Geografia e as ciências afins e contribuir com posteriores diagnósticos que envolvam a dinâmica climática numa bacia usaram o método de análise de agrupamento de Ward para mapear a pluviosidade homogênea na bacia hidrográfica do rio Goiana-PE e posteriormente verificar os principais fatores físicos, internos à bacia, que influenciam na espacialização da chuva, empregando a distância euclidiana como medida de dissimilaridade. Os autores obtiveram resultados que mostraram que a bacia hidrográfica do rio Goiana possui três regiões pluviometricamente homogêneas e que estão fortemente influenciadas pelos fatores marítimos (longitudinais) e topográficos.

Andre *et al.* (2008) visando dar uma contribuição ao estudo das condições climáticas do Estado do Rio de Janeiro, e uma melhor Classificação Climática por meio da identificação de regiões homogêneas em precipitação, utilizou médias mensais da precipitação de 48 estações meteorológicas, em um período de 30 anos (1971-2000). A análise hierárquica de agrupamento, a orografia e a proximidade do mar, mostraram que o Estado do Rio de Janeiro pode ser dividido, quanto à precipitação, em seis regiões pluviometricamente homogêneas o

que possibilitou classificar as estações meteorológicas pelo método de classificação não hierárquica *k-means*. A região norte do Estado, com precipitações anuais em torno de 870 mm é a mais seca, e a região da encosta sul da Serra do Mar, com 2020 mm, é a mais chuvosa. Mas, em ambas as regiões, os valores da precipitação da estação chuvosa representam em torno de 70% dos totais anuais.

## MATERIAL E MÉTODOS

### Dados Meteorológicos

Para este estudo, parte de um projeto mais amplo, foram utilizados a título de demonstração, dados de 522 pontos distribuídos pelo Nordeste do Brasil (NEB) (Figura 1) interpolados pelo método "Krigagem", com base nos dados do National Climatic Data Center (NCDC), obtidas de séries médias mensais de precipitação de toda a região Nordeste, gradeadas de 0,5 em 0,5 grau durante um período de 30 anos (1961-1990), disponível em [www.ncdc.noaa.gov](http://www.ncdc.noaa.gov).

Neste trabalho, uma técnica hierárquica de agrupamento (Método de Ward) é usada para dividir a área do Nordeste do Brasil em um número limitado de zonas climaticamente homogêneas baseadas na variável meteorológica da precipitação mensal para 522 pontos amostrais gradeados; além disso, é aplicada a análise de componentes principais (ACP), possivelmente a técnica multivariada estatística mais amplamente usada em ciências atmosféricas.

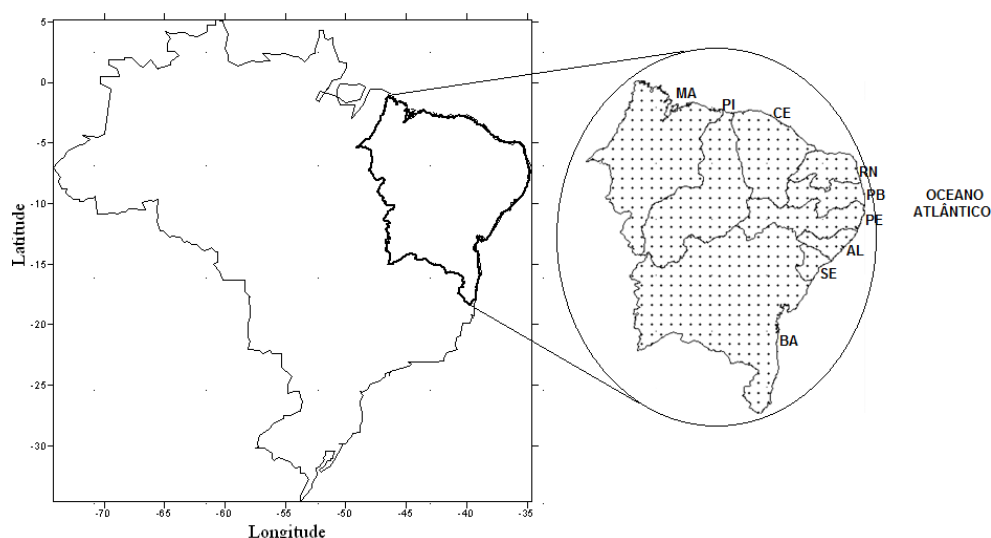


Figura 1. Distribuição espacial dos 522 pontos amostrais na região Nordeste do Brasil.

## PROCEDIMENTOS TÉCNICO-OPERACIONAIS

### Aplicação da Análise de Componentes Principais

A metodologia fundamental deste estudo baseou-se na técnica de Análise Multivariada. Considerou-se a Análise Fatorial (AF) em Componentes principais (CP) nos modos espacial e temporal para definir os padrões das precipitações na Região Nordeste do Brasil (NEB).

A análise de componentes principais é essencialmente um método de redução de dados que objetiva produzir um número pequeno de variáveis derivadas que possam ser usadas no lugar do número maior de variáveis originais para simplificar análises subsequentes dos dados. Este método permite a redução da dimensionalidade dos pontos representativos das amostras, pois,

embora a informação estatística presente nas  $n$ -variáveis originais seja a mesma das  $n$ -componentes principais, é comum obter em apenas 2 ou 3 das primeiras componentes principais mais que 90% desta informação.

As componentes principais  $y_1, y_2, y_3, \dots, y_q$  são definidas por combinações lineares das variáveis originais  $x_1, x_2, x_3, \dots, x_q$  que são não-correlacionadas e quantificam as máximas proporções de variação nos dados originais, isto é,  $y_1$  conta para o máximo de variância de todas as possíveis combinações lineares de  $x_1, x_2, x_3, \dots, x_q$ ,  $y_2$  conta para o máximo de variância não-explicada por  $y_1$ , e assim por diante.

Explicitamente, as componentes principais são obtidas a partir de  $x_1, x_2, x_3, \dots, x_q$  assim:

$$y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1q}x_q$$

$$y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2q}x_q$$

$\vdots$

$$y_q = a_{q1}x_1 + a_{q2}x_2 + \dots + a_{qq}x_q$$

em que os coeficientes  $a_{ij}$  ( $i = 1, \dots, q; j = 1, \dots, q$ ) sejam escolhidos de tal forma que retenham a máxima variância e sejam não-correlacionados.

Os coeficientes que definem as componentes principais são dados pelo que são conhecidos como autovetores da matriz de covariância  $S$ , ou matriz de correlação,  $R$ . As componentes derivadas a partir de  $S$  devem diferir consideravelmente daquelas derivadas de  $R$ , e não há necessariamente qualquer relação entre elas.

Se as primeiras variáveis derivadas (as componentes principais) quantificarem uma boa proporção da variância total das variáveis observadas, elas podem ser usadas tanto para fornecer um resumo conveniente dos dados quanto para simplificar análises subsequentes.

Os coeficientes definidos pelas componentes principais são soluções de uma série de equações envolvendo os elementos da matriz de covariância observada, embora quando as variáveis originais estão em escalas muito diferentes, é aconselhável extraí-los a partir da matriz de correlação; foi o caso desta pesquisa.

A escolha do número de componentes que são adequadas para resumir um conjunto de dados multivariados é geralmente baseado num relativo número de procedimentos:

- Reter apenas as componentes que explicam porcentagens consideráveis da variação total das variáveis originais. Valores entre 70 e 90% são usualmente sugeridos, embora valores pequenos possam ser apropriados como número de variáveis,  $q$ , ou número de sujeitos,  $n$ ;
- Excluir aquelas componentes principais com variâncias menores que a média. Quando as componentes são extraídas a partir da matriz de correlação observada, isto implica excluir componentes com variância menor que um. (Esta é uma aproximação muito popular, no entanto, tem sido criticada; veja, por exemplo, Preacher e MacCallam, 2003.);
- Plote das componentes (*scree plot*).

### **Aplicação da Análise de Agrupamento**

Utilizou-se a Análise de Agrupamentos (AA) para classificar os pontos pluviométricos (não postos pluviométricos) do NEB quanto ao comportamento espacial semelhante.

A classificação é um importante componente em toda pesquisa científica. Existem essencialmente duas técnicas estatísticas interessadas na classificação. A análise de agrupamento (EVERITT, LANDAU, e LEESE, 2001) que objetiva descobrir grupos de

observações a partir de dados inicialmente não-classificados. A análise de função discriminante (EVERITT e DUNN, 2001) que trabalha com dados que já estão classificados dentro de grupos. Na análise de agrupamento existem dois tipos de técnicas principais: divisiva e hierárquica (KAUFMAN e ROUSSEAUW, 1990). O objetivo da técnica divisiva é separar um conjunto de objetos dentro de grupos consistentes, enquanto na hierárquica, os objetos são progressivamente agregados até eles estarem unidos em um único agrupamento.

A técnica de agrupamento hierárquico interliga as amostras por suas associações, produzindo um dendrograma onde as amostras semelhantes, segundo as variáveis escolhidas, são agrupadas entre si. Os dendrogramas são especialmente úteis na visualização de semelhanças entre amostras ou objetos representados por pontos em espaço com dimensão maior do que três, onde a representação de gráficos convencionais não é possível.

Para agrupar as regiões com média da precipitação mensal similar, aplicou-se o método hierárquico aglomerativo de Ward, tendo como medida de dissimilaridade a distância euclidiana (EVERITT e DUNN, 1991):

$$d_e = \left[ \sum_{j=1}^n (P_{p,j} - P_{k,j})^2 \right]^{0,5}$$

Onde  $d_e$  é a distância euclidiana; e  $P_{p,j}$  e  $P_{k,j}$  são as variáveis quantitativas  $j$  dos indivíduos  $p$  e  $k$ , respectivamente.

O método de Ward forma grupos, minimizando a dissimilaridade, ou minimizando o total das somas de quadrados dentro de grupos, também conhecida como soma de quadrados dos desvios ( $SQD$ ). Em cada etapa do procedimento, são formados grupos, de tal maneira que a solução resultante tenha o menor  $SQD$  dentro de grupos. Nessas etapas, são consideradas as uniões de todos os possíveis pares de grupos e, os dois que resultam em menor aumento de  $SQD$  são agrupados até que todos os grupos formem um único, reunindo todos os indivíduos (EVERITT e DUNN, 1991).

## RESULTADOS E DISCUSSÃO

### Padrões espaciais obtidos usando ACP

Na Figura 2 é apresentada a distribuição espacial dos totais anuais médios de precipitação no Nordeste do Brasil (NEB) pelo método de interpolação krigagem. A partir da figura constata-se que a região do extremo noroeste caracteriza-se como a mais chuvosa com valores acima de 2000 mm e a região menos chuvosa encontra-se na parte central se estendendo a leste (Raso da Catarina - BA e sertão e agreste de PB e PE) com valores inferiores a 600 mm.

A média mensal das precipitações em mm para o período de 1961-1990 está representada na Figura 3, a qual mostra que o mês mais chuvoso é março com valor acima de 180 mm e o menos chuvoso é agosto com valor inferior a 20 mm.

Os resultados das Figuras 2 e 3 são semelhantes aos encontrados nos atlas pluviométricos das agências ANA (disponível em [www.ana.gov.br](http://www.ana.gov.br)) e FUNCEME (disponível em [www.funceme.br](http://www.funceme.br)).

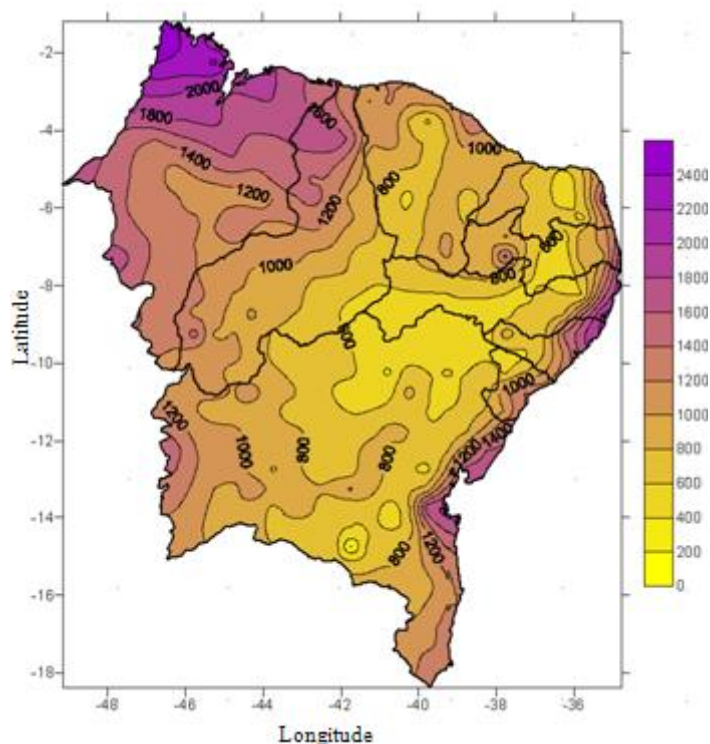


Figura 2. Distribuição espacial dos totais médios anuais em mm da precipitação pluviométrica no Nordeste do Brasil com base nos dados do NCDC.

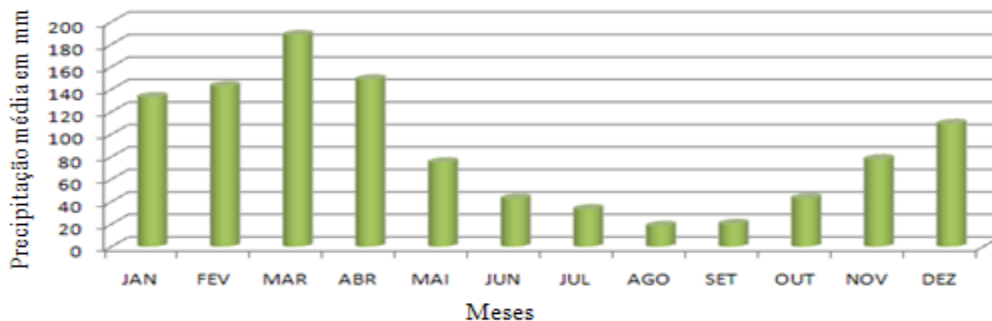


Figura 3 – Precipitação média mensal em mm para o Nordeste do Brasil com base no NCDC.

Após aplicado o método de Análise de Componentes Principais, reteve-se três primeiras cargas fatoriais comuns temporais rotacionadas, que explicaram 34,95%, 30,76% e 30,45% da variância dos dados, respectivamente, totalizando mais que 96% de variância explicada (Tabela 1). O quarto fator apresentou 1,4% de variância. Ele foi rejeitado devido ser recomendável descartar cargas cujas variâncias sejam inferiores a 5% (O'LENIC e LIVEZEY, 1998), tal autovalor é inferior a 1 (0,16) pelo critério de Kaiser (GARAYALDE *et al.*, 1986), e efetuando-se uma pequena inspeção no "scree plot" (Figura 4) o gráfico apresenta tendência de paralelismo com o eixo das abscissas a partir do quarto fator, ou seja, sem influência.

Tabela 1

**Autovalores em ordem decrescente e a contribuição (%) à variância total dos dados médios mensais rotacionados da precipitação pluvial no NEB.**

Fator	Autovalor	Variância Explicada (%)	Variância Acumulada (%)
1	4,19	34,95	34,95
2	3,69	30,76	65,70
3	3,66	30,45	96,17
4	0,16	1,30	97,46
...	...	...	...
12	0	0	100

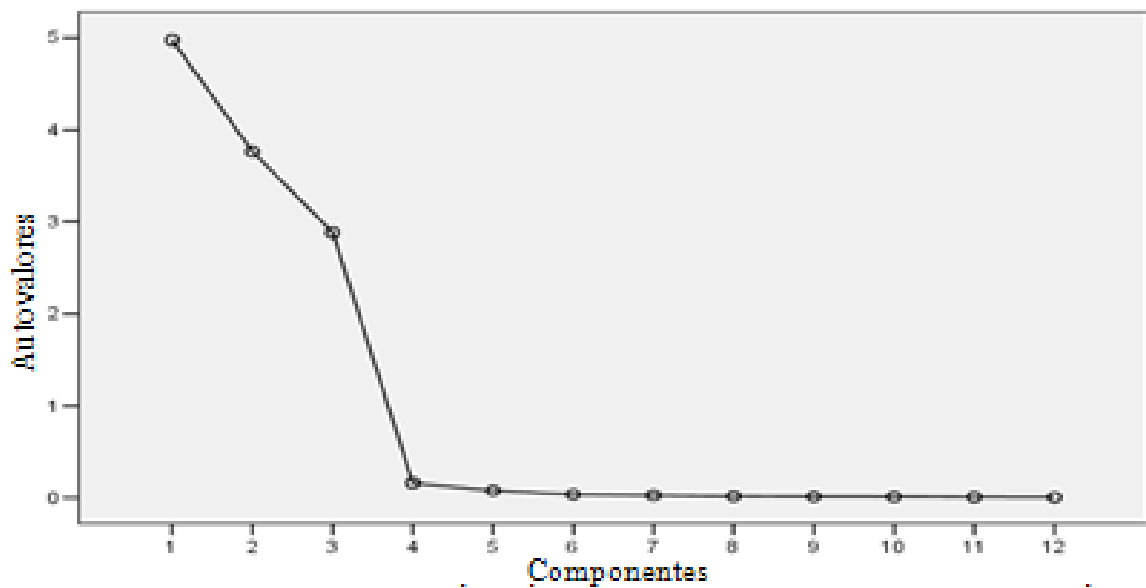


Figura 4 - Scree plot das componentes versus autovalores



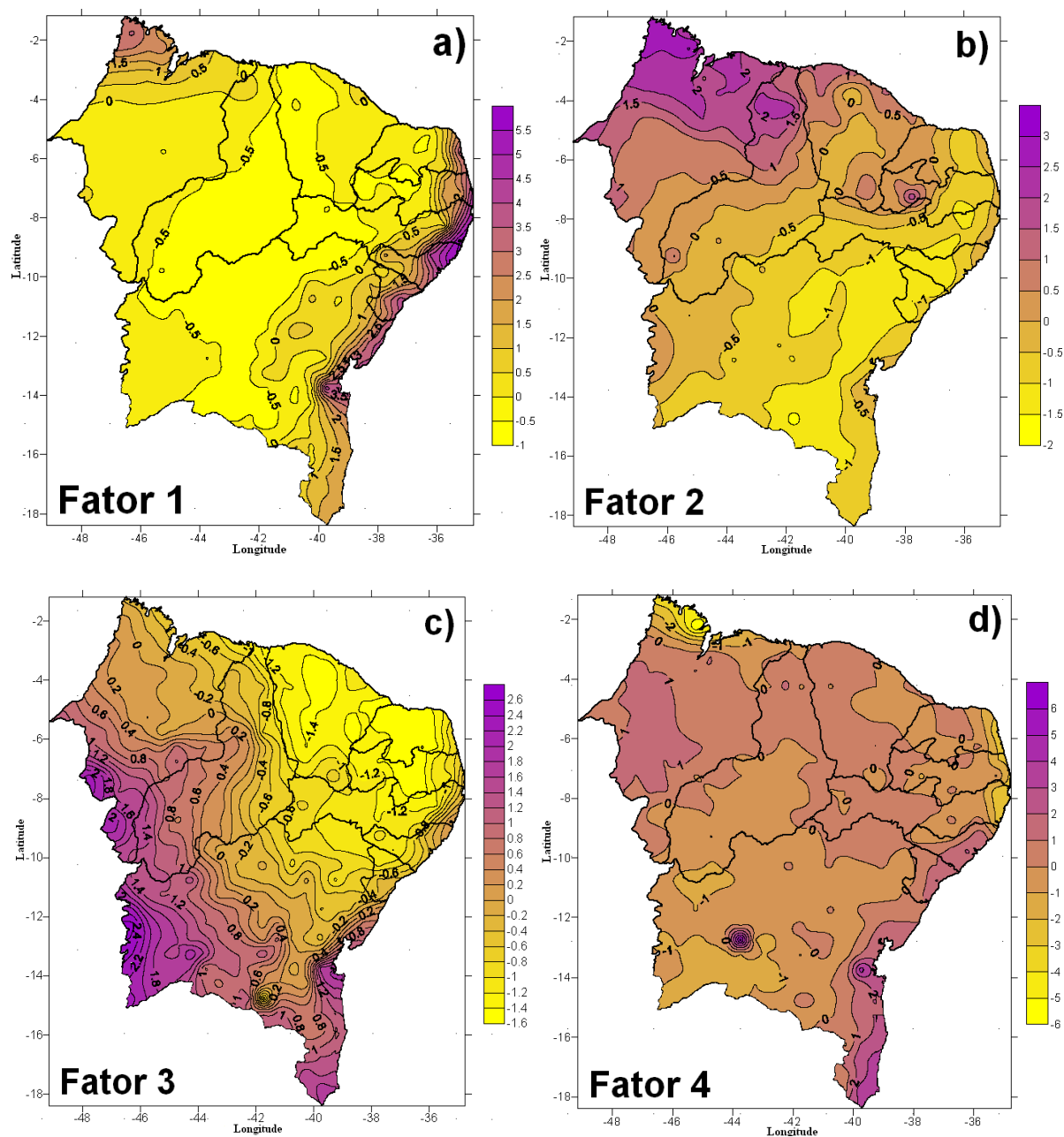


Figura 5 – Distribuição das cargas (pesos) médias dos fatores de precipitação, Fator 1 (a), Fator 2 (b), Fator 3 (c) e Fator 4 (recusado) (d) no período de 1961 a 1990.

O primeiro fator, que explica 34,95% da variância, apresenta forte correlação anual com máximo entre os meses de junho a setembro e mínimo entre os meses de novembro a março (Figura 6) e sua configuração espacial, os escores, mostra contrastes zonais positivos (alcançando valores acima de 5,0) (Figura 5a) principalmente na parte do litoral, por toda a costa leste, dando o indicativo de uma forte influência do sistema de distúrbios de leste que atuam desde o norte do Rio Grande do Norte até a Bahia e as contribuições para intensificação associadas às circulações de brisas marítimas e terrestres. Ainda podem ser encontrados valores positivos no extremo norte do Maranhão em que a influência pela proximidade com a floresta amazônica é manifestada (Figura 5a).

As correlações entre as cargas fatoriais e as séries mensais dos pontos de dados amostrais, associadas aos 3 fatores são mostradas na Figura 6.

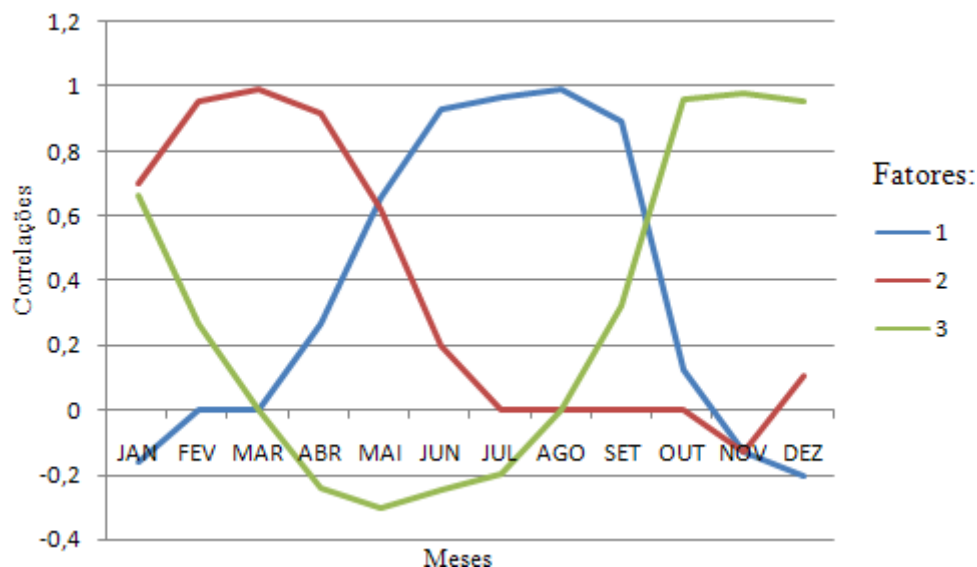


Figura 6 – Correlações dos três primeiros fatores rotacionados associados aos meses do ano.

O segundo fator explica 30,76% da variância total dos dados e tem correlações maiores que 0,8 nos meses de fevereiro a abril (Figura 6). Na distribuição espacial referente a este fator apresenta-se cargas fatoriais bastante distintas. Podem ser identificados núcleos negativos em praticamente toda a porção sul e agreste da região com mínimo de junho a novembro (Figura 6), e contribuições positivas com máximos de fevereiro a maio notadamente no estado do Maranhão e norte do estado do Piauí (Figura 5b).

Por fim, o terceiro fator comum explicando 30,45% da variância apresenta altas correlações superiores a 0,8 nos meses de outubro a dezembro (Figura 6). A distribuição espacial referente a este fator mostra núcleos negativos da região do Raso da Catarina na Bahia até o extremo norte, enquanto na parte sul e oeste de toda a área de estudo se observa núcleos positivos chegando a 2.6 (Figura 5c). Esse período de precipitação é influenciado pela influência dos sistemas frontais que é um importante sistema produtor de precipitação, atuando basicamente na região sul do NEB. Atuam nos meses de novembro a fevereiro, principalmente na parte sul da Bahia, decrescendo para o norte (KOUSKY, 1979).

A Figura 7 mostra a climatologia mensal dos meses associados ao 1º fator, em que o mês de junho apresenta valores de precipitação até 300 mm, o mês de julho até 320 mm, o mês de agosto até 170 mm, o mês de setembro até 130 mm. Percebe-se a semelhança entre a representação do Fator 1 (Figura 5a) e a Figura 7 (a, b, c e d, respectivamente).

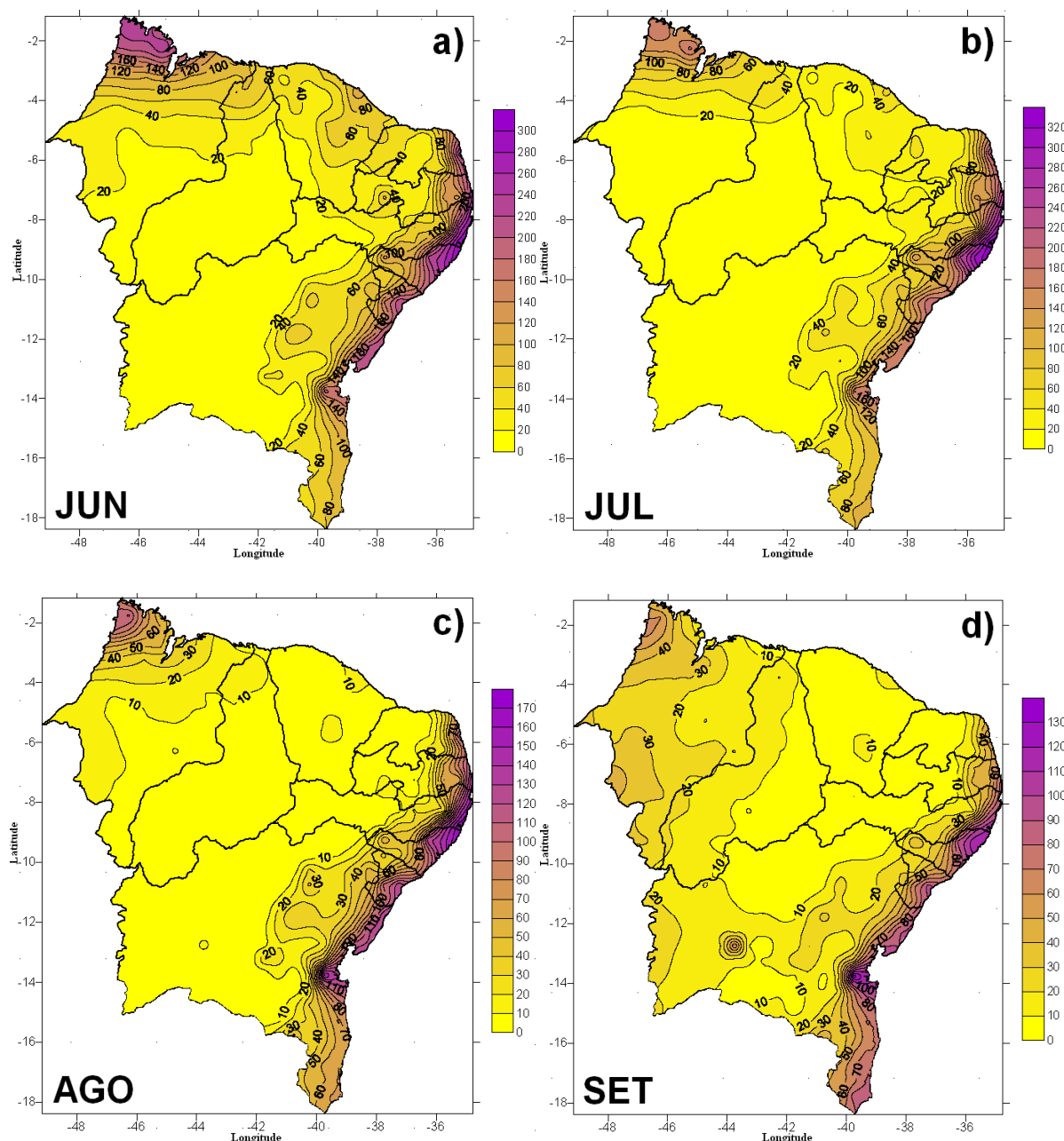


Figura 7 – Distribuição dos padrões médios mensais de precipitação em mm, para os meses de Junho (a), Julho (b), Agosto (c) e Setembro (d) no período de 1961 a 1990.

Os pesos referentes ao 2º fator estão representados na Figura 8, cujos máximos e intensidades estão na região norte associados à zona de convergência intertropical (ZCIT) e aos VCAS, e diminuem até o mês de maio quando se concentram mais ao litoral baiano e norte maranhense.

A Figura 8 mostra a climatologia mensal dos meses associados com o 2º fator, em que o mês de fevereiro apresenta valores de precipitação até 340 mm, o mês de março até 440 mm, o mês de abril até 450 mm, o mês de maio até 340 mm. Percebe-se a semelhança entre a representação do Fator 2 (Figura 5b) e a Figura 8 (a, b, c e d, respectivamente).

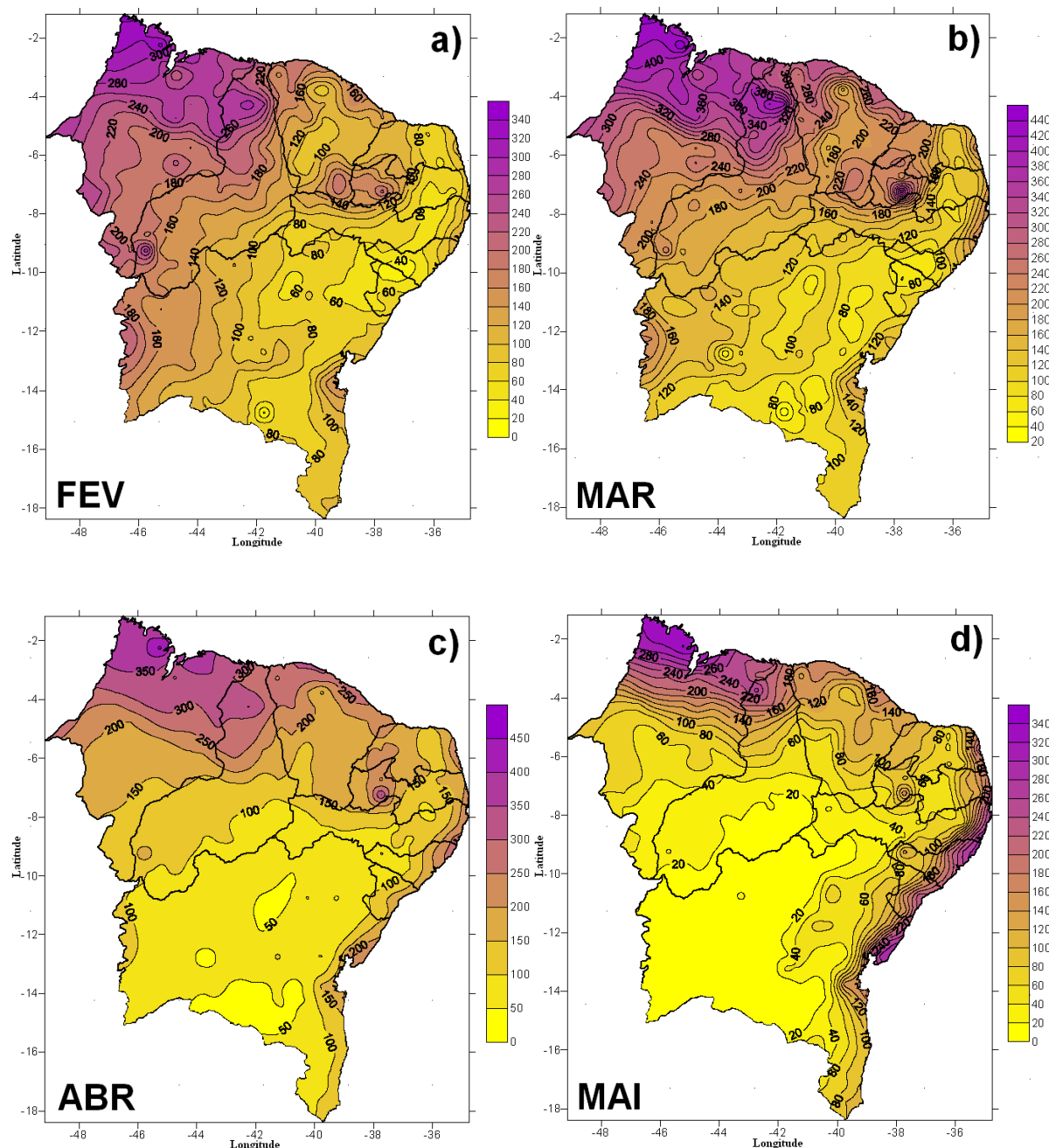


Figura 8 – Distribuição dos padrões médios mensais de precipitação em mm, para os meses de Fevereiro (a), Março (b), Abril (c) e Maio (d) no período de 1961 a 1990.

A Figura 9 mostra a climatologia mensal dos meses associados com o 3º fator, em que o mês de outubro apresenta valores de precipitação até 140 mm, o mês de novembro até 220 mm, o mês de dezembro até 260 mm, o mês de janeiro até 260 mm. A semelhança entre a representação do Fator 3 (Figura 5c) e a Figura 9 (a, b, c e d, respectivamente), é notória quando se compara as respectivas Figuras.

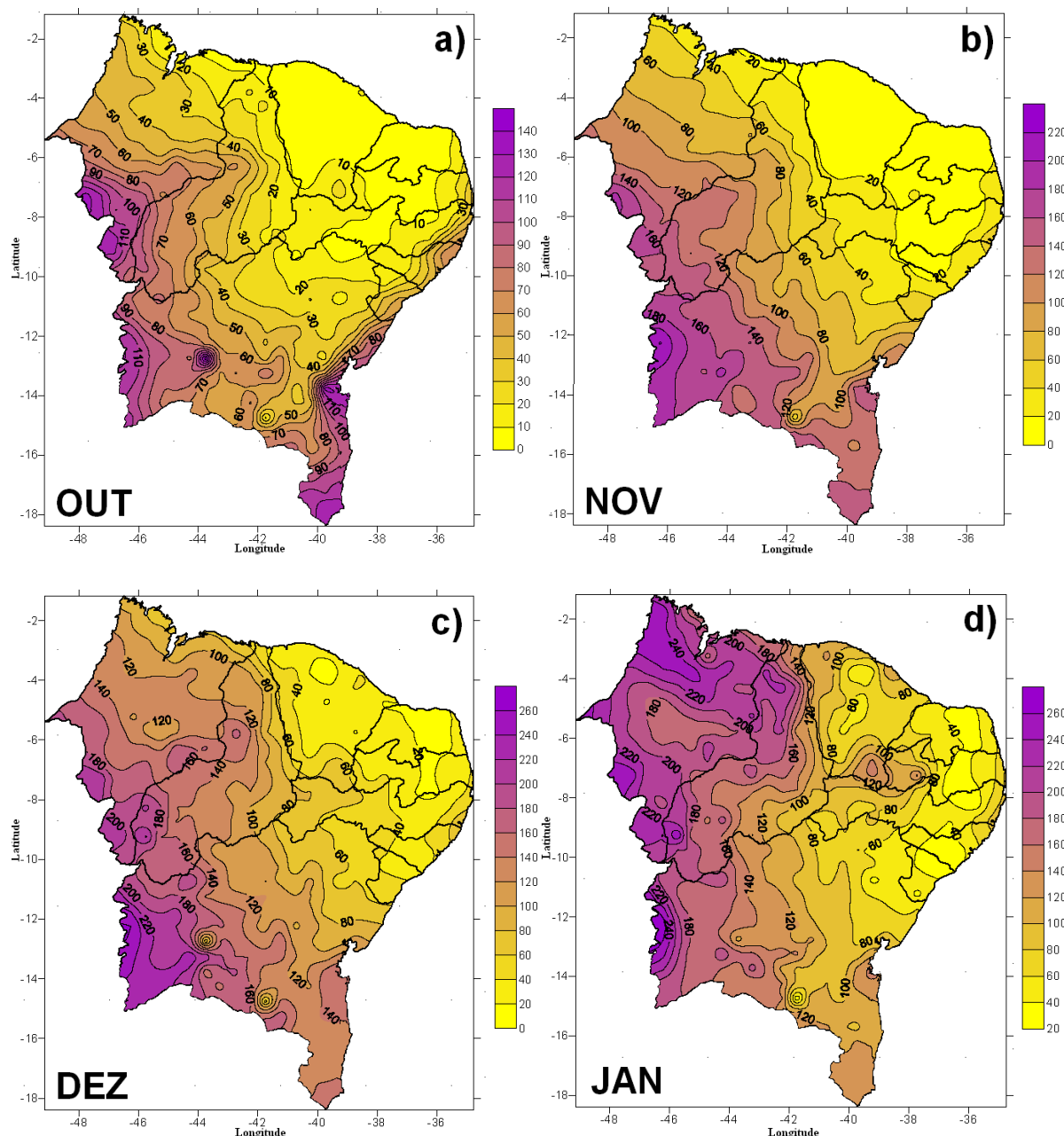


Figura 9 – Distribuição dos padrões médios mensais de precipitação em mm, para os meses de Outubro (a), Novembro (b), Dezembro (c) e Janeiro (d) no período de 1961 a 1990.

#### Caracterização homogênea usando AA

Como variáveis classificatórias foram utilizadas as médias mensais e anuais da precipitação, além das respectivas coordenadas geográficas de cada ponto pluviométrico.

A distribuição de chuvas por todo o nordeste brasileiro, decrescentes do litoral para o interior, sugere a identificação de regiões homogêneas quanto ao regime pluvial. Foram determinadas quatro regiões homogêneas (Figura 12), denominadas de Grupo, a partir do agrupamento dos três fatores comuns rotacionados de pluviosidade utilizando o método de agrupamento hierárquico aglomerativo de Ward.

Através do gráfico da inércia (Figura 10) foi possível estabelecer "a região do corte" no dendrograma (Figura 11) e consequentemente se ter o número de 4 grupos:

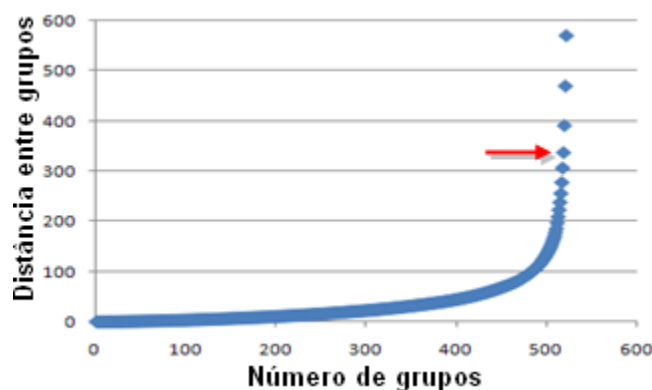


Figura 10 – Gráfico representativo da Curva de Inércia.

É observado através da Figura 10 uma seta indicando onde, a partir da qual, há um salto nos valores, ou seja, que a partir desse ponto em destaque, as distâncias entre os grupos tornam-se consideráveis, daí a justificativa para a escolha de quatro grupos.

Desta forma, usando este critério, chegou-se a um dendrograma (diagrama de árvore) resultante do agrupamento das componentes da precipitação dos 522 pontos de grade analisados (não exibido).

As regiões estão representadas na Figura 11. Nota-se um padrão claro na distribuição da precipitação, com uma orientação no sentido leste-oeste das regiões homogêneas a qual pode indicar a influência da maritimidade e continentalidade sobre a distribuição da precipitação na região nordeste do Brasil. No grupo 1 encontram-se os pontos pluviométricos localizados na parte extremo noroeste, notadamente norte dos estados do Maranhão e Piauí, e regiões adjacentes.

O grupo 2 inclui a região central e parte da região norte e sul (semiárido), enquanto no grupo 3 encontram-se os pontos pluviométricos representativos do Oeste em direção ao cerrado. No grupo 4 estão inseridos os pontos pluviométricos localizados no extremo-leste, contornos litorâneos.

Observa-se pela distribuição espacial do primeiro fator (Figura 5a), que as maiores contribuições (pesos) com valores superiores a 5, encontram-se na região homogênea 4 (em tom vermelho) (Figura 11). A distribuição deste primeiro fator relaciona às chuvas de junho a setembro (Figura 7) e está associado às precipitações decorrentes dos distúrbios de leste (sistemas de brisa e ventos alísios de sudeste) (KOUSKY, 1980).

O padrão espacial associado ao segundo fator (Figura 5b) apresenta valores superiores a 2,0 na região homogênea 1 (em tom amarelo) (Figura 11). A distribuição das chuvas, deste segundo fator, nos meses de fevereiro a maio (Figura 8) pode estar associado aos vórtices ciclônicos de ar superior (VCAS) (KOUSKY, 1979).

A configuração espacial do terceiro fator apresenta as maiores contribuições da precipitação, com valores superiores a 2,0, na região homogênea 3 (em tom azul) (Figura 11), estando associadas com a atuação da entrada de frentes sobretudo nos meses de outubro a janeiro (Figura 9), enquanto os valores inferiores a 0 foram encontrados na região homogênea 2 (em tom de verde) (Figura 11) cujo principal sistema indutor de precipitação é a durante a fase da Zona de Convergência Intertropical (ZCIT) (UVO, 1989).



Legenda:

- amarelo (Grupo 1)
- verde (Grupo 2)
- azul (Grupo 3)
- vermelho (Grupo 4)

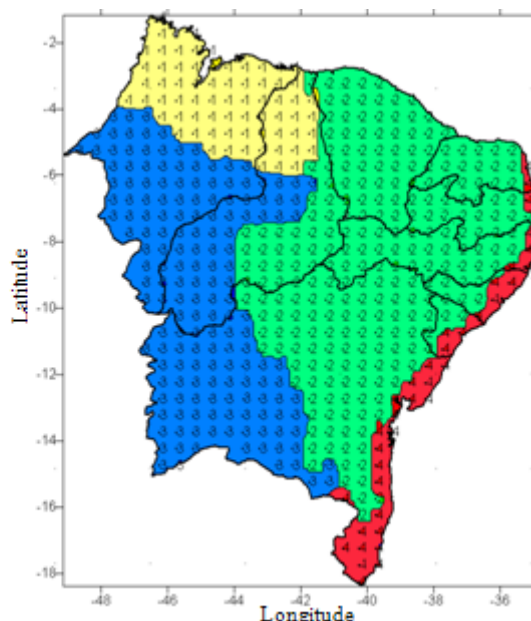


Figura 11 – Distribuição das regiões pluviometricamente homogêneas da Região Nordeste

## CONCLUSÃO

Com relação a ACP, foi uma ferramenta imprescindível para realização destas observações climatológicas, pois através da mesma foi possível identificar os grupos de variáveis intercorrelacionadas, reduzindo o número de variáveis a serem estudadas e reescrevendo uma nova série de dados.

O primeiro fator comum das precipitações teve correlações acima de 0,8 nos meses de junho-setembro e está associado aos sistemas ou distúrbios de leste, maior disponibilidade de umidade devido à proximidade do oceano e circulações de brisas marítimas e terrestres. O segundo fator apresenta altas correlações nos meses de fevereiro-abril associados a ZCIT e aos VCAS. Enquanto o terceiro fator apresenta correlações significativas nos meses de outubro-dezembro, em decorrência dos sistemas frontais.

A classificação da precipitação pluvial usando o método aglomerativo de Ward segundo o critério da inércia possibilitou a obtenção de quatro sub-regiões (grupos) homogêneas no Nordeste do Brasil.

O peso do primeiro fator comum mostrou que as maiores contribuições ocorreram na parte litorânea da região nordeste (Grupo 4). O peso associado ao segundo fator apresentou maiores contribuições na região extremo noroeste (Grupo 1). O peso do terceiro fator evidenciou o regime pluviométrico na região sul e oeste (Grupo 3), e finalmente para estes três fatores, a região agreste (Grupo 2) apresentou as menores contribuições.

A Análise Fatorial das séries temporais representadas pelos padrões de precipitação possibilitaram uma melhor compreensão dos aspectos físicos responsáveis pela variabilidade mensal da área estudada.

## REFERÊNCIAS

- .AHMED, B. Y. M.: Climatic classification of Saudi Arabia: an application of factor-cluster analysis, **GeoJournal**, 41.1, 69–84, 1997.
- .ANDRADE, E. M., SILVEIRA, S. S. & AZEVEDO, B. M. Investigação da estrutura multivariada da evapotranspiração na região Centro Sul do Ceará pela Análise de Componentes Principais. **Revista Brasileira de Recursos Hídricos**, vol. 8, n. 1, 2003.
- .ANDRE, R. G. B, MARQUES, V. da S., PINHEIRO, F. M. A, FERRAUDO, A. S. Identificação de regiões pluviometricamente homogêneas no estado do Rio de Janeiro, utilizando-se valores mensais. **Revista Brasileira de Meteorologia**, v.23, n.4, 501 - 509, 2008.
- .ASSIS, F. Neto. De; Arruda, H. V. de; Pereira, A. R. **Aplicação de estatística à climatologia - Teoria e pratica**. Editora e Gráfica Universitária (UFPEL), 1996. 161p.
- .BRAGA, C.C., de MELO, M.L.D., MELO, E.C.S. Análise de agrupamento aplicada a distribuição da precipitação no Estado da Bahia. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 10, Brasília-DF. **Anais...** Sociedade Brasileira de Meteorologia. p.1857-62, 1998.
- .DEGAETANO, A. T. and SHULMAN, M. D.: A climatic classification of plant hardiness in the United States and Canada, **Agric. For. Meteorol.**, 51, 333–351, 1990.
- .DEGAETANO, A. T.: Spatial grouping of United States climates stations using a hybrid clustering approach, **Int. J. Climatol.**, 21, 791–807, 2001.
- .DINIZ, M. de C. S. **Variabilidade climática e ocorrência de cheias na zona semi-árida da bacia hidrográfica do São Francisco**. Dissertação de mestrado em Meteorologia, UFCG (Universidade Federal de Campina Grande), Campina Grande, PB, 2007.
- .EVERITT, B.S.; DER, G. A. **Handbook of statistical analyses using SAS**. London: Chapman & Hall, 1977.
- .EVERITT, B.S.; DUNN, G. **Applied multivariate analysis**. London: Edward Arnold, 1991. 400p.
- .EVERITT, B. S., Landau, S., and Leese, M. **Cluster Analysis** (4th ed). London: Arnold, 2001.
- .EVERITT, B. S. and Dunn, G. **Applied Multivariate Data Analysis** (2nd ed). London: Arnold, 2001.
- .GARAYALDE, E.J.G.; da SILVA, M.G.A.; TAVARES, A. de SÁ. Classificação meso-climática da região sul do Brasil pela análise de Componentes Principais. In: CONGRESSO INTERAMERICANO DE METEOROLOGIA, 1., CONGRESSO BRASILEIRO DE METEOROLOGIA, 1986, Brasília. **Anais...**,Brasília: Sociedade Brasileira de Meteorologia, v. 1, 1986. p. 119-124.
- .GONG, X. and RICHMAN, M. B.: On the application of cluster analysis to growing season precipitation data in North America East of the Rockies, **J. Clim.**, 8, 897–931, 1995.
- .HALKIDI, M., BATISTAKIS, Y., and VAZIRGIANNIS, M.: On clustering validation techniques, **Journal of Intelligent Information Systems**, 17:2/3, 107–145, 2001.
- .KAUFMAN, L. and ROUSSEUW, P. J.: **Finding Groups in Data: An Introduction to Cluster Analysis**, Wiley, New York, 1990.
- .KOUSKY, V. E. Frontal influences on northeast Brazil. **Monthly Weather Review**, v.107, n.17, p.1140-1153, 1979.
- .KOUSKY, V. E. Diurnal rainfall variation in Northeast Brazil. **Monthly Weather Review**, v. 108, n.4, p.488-498, 1980.
- .MALDONADO, F. D. et al. Caracterización de câmbios e influencia climática sobre la cobertura vegetal em el semiárido brasileiro mediante Análisis de Componentes Principales. **Revista de Teledetección**, 13: 27-36, 2000.
- .O'LENIC, E. A. and LIVEZEY, R. E. Practical considerations in the use of rotated principal component analysis (RPCA) in diagnostic studies of upper-air height fields, **Mon. Weather Rev.**, 116, 1682–1689, 1998.
- .PRATES, J. E. **Controles Associados à Distribuição Espacial de Precipitação no Verão em Minas Gerais: Aspectos Fisiográficos e Meteorológicos**. São Paulo, 1994. 178p. Tese (Doutorado em Meteorologia) - Universidade de São Paulo.



- .PREACHER, K. J. and MacCallum, R. C. (2003) Repairing Tom Swift's Electric Factor Analysis. **Understanding Statistics**, 2, 13-44.
- .RAMOS, M. C.: Divisive and hierarchical clustering techniques to analyse variability of rainfall distribution patterns in a Mediterranean region, **Atmos. Res.**, 57, 123-138, 2001.
- .SANTOS, A. M. dos, GALVÍCIO, J. D., MOURA, M. S. B. de. Homogeneização da precipitação pluviométrica na Bacia Hidrográfica do Rio Goiana-PE, com método de análise de agrupamento. **RBGF – Revista Brasileira de Geografia Física** Recife-PE v. 1, n. 1, Mai/Ago 2008, p.14-27.
- .SILVA, A. A. G. da; BRAGA, C. C.; OLIVEIRA, L. S. Análise de Agrupamento Aplicada ao Regime Pluviométrico da Costa Leste e Norte do Nordeste do Brasil. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 8, Porto Alegre-RS. **Anais...** Sociedade Brasileira de Meteorologia. p.1758-63, 1996.
- .SILVA, F. A.; LERMEN, V. k. ; NERY, J. T.; Variabilidade interanual da precipitação na bacia do rio Iguaçu. **Revista da Universidade Estadual de Maringá** – Paraná / Brasil. Páginas 1439 a 1444, ano 2001.
- .SILVA, H. D. da S. J. **Evapotranspiração na Região Nordeste do Brasil utilizando Estatística Multivariada**. Dissertação de mestrado em Meteorologia, UFCG (Universidade Federal de Campina Grande), Campina Grande, PB, 2004.
- .UVO, C. R. B. **The Intertropical Convergence Zone and its relationship with the precipitation over north-northeast region of Brazil**. Tech. Rep. INPE-4887-TDL/378, Available from Instituto Nacional de Pesquisas Espaciais, 12200, São José dos campos, SP, Brazil, 1989, 82p.
- . Ward, J. H. (1963). **Hierarchical grouping to optimize an objective function**. *J. Am. Statist. Assoc.* 58, 236-244